

LOW-LEVEL RADIOLOGICAL WASTE EVALUATION ASSOCIATED WITH VARIOUS BASE REALIGNMENT AND CLOSURE ACTIVITIES

January 13, 2011

Prepared by

Kurt Picel and Robert Johnson
Environmental Science Division
Argonne National Laboratory

Under contract to Battelle Memorial Institute

Disclaimer

This report was prepared as an account of work sponsored by Battelle Memorial Institute. Neither the United States Government nor any agency thereof, nor UChicago Argonne, LLC, nor any of their employees or officers, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of document authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof, Argonne National Laboratory, or UChicago Argonne, LLC.

EXECUTIVE SUMMARY

The objective of the current study was to review the remedial action program for the soil associated with the sewer system at Hunters Point Shipyard in San Francisco, California, for the purpose of reducing the cost of low-level radiological waste (LLRW) identification and disposal. The review resulted in the following suggested improvements, expected results, and suggested implementation steps:

- Revamp the sodium iodide (NaI) pad screening method to one that effectively identifies locations with truly elevated radium-226 (Ra-226) levels. Expected results would include:
 - Reduced number of biased samples that do not find contamination and reduced overall sampling load, reduced false identification of LLRW due to analytical uncertainty in biased samples, and improved positive identification of elevated areas. Implementation would:
 - Use pad-specific background levels to establish a gross activity investigation level.
 - Confirm that flagged areas are actually elevated before proceeding with full biased sampling; analyze highest areas first in on-site laboratory.
 - Determine the gross activity response of the NaI detector to 1 pico-Curie per gram (pCi/g) of Ra-226 to provide a basis for investigation levels.
 - Alternatively, deploy a multichannel analyzer and make specific Ra-226 activity concentration measurements on pads.
- Implement improvements in the on-site laboratory equipment and methods to reduce the uncertainty in Ra-226 measurements and reduce the minimum detectable concentration (MDC) to well below action levels. Expected results would include:
 - Reduced false identification of LLRW due to high relative analytical uncertainty and high bias. Implementation would:
 - Analyze Ra-226 on the basis of bismuth (Bi-214), accounting for Ra-222 disequilibrium by allowing for ingrowth or by applying a correction factor.
- Expand the background dataset to encompass full Ra-226 variability in all soil types being remediated. Expected results would include:
 - Reduced false identification of LLRW due to under representation of background variability. Implementation would:
 - Review existing pad datasets to identify uncontaminated pads or portions of pads and establish a larger background dataset from existing data that encompasses the full range of Ra-226 background.

- Implement the 1-pCi/g plus background requirement using the upper end of the background data distribution rather than the mean of the distribution while retaining the 2-pCi/g upper limit. Expected results would include:
 - Reduced false identification of LLRW for soil with background levels falling in the upper half of the background distribution. Implementation would:
 - Determine the 95th percentile of the normal distribution of a large, robust, background dataset that encompasses all soil conditions and apply the 1-pCi/g Ra-226 criterion to this threshold.
 - Alternatively, apply the 1-pCi/g Ra-226 criterion to the 95% upper confidence interval on the mean background, which at least would account for uncertainty in what true mean background activity concentrations are for the site.
- Use Ra-228 levels to evaluate the actual presence of Ra-226 contamination in areas where no clearly elevated levels of Ra-226 are present. Expected results would include:
 - Reduced false identification of LLRW through a determination that contamination is not present. Implementation would:
 - Determine the Ra-226 to Ra-228 ratio and ratio variability in a large set of existing soil results at background levels and use this information to determine a threshold that would identify when Ra-226 was clearly elevated above background conditions.
 - Use this ratio as part of a body of evidence for making determinations that Ra-226 is not actually present.
- Revamp pad sampling protocols: implement the criterion as a wide-area average requirement and related elevated area comparison values using Multi-Agency Radiation Survey and Site Investigation Manual (MARSSIM) principles. Expected results would include:
 - Reduced identification of LLRW by applying the Ra-226 criterion over a defined volume of soil as an average, reduced sampling costs, and reduced false identification of LLRW due to analytical uncertainty for single samples. Implementation would:
 - Implement pad protocols based on MARSSIM principles: select an appropriate parametric or nonparametric statistical test and design pad sampling accordingly, and implement a wide-area criterion and an elevated measurement comparison standard for small elevated areas.
 - Continue to implement a full pad scan to identify elevated areas, but with improvements as suggested above.
 - Confirm on-site laboratory results barely above cleanup levels when pads are not clearly contaminated.
 - Eliminate repeat systematic sampling on pads when pad remediation is required.

1 INTRODUCTION

On behalf of the United States Department of the Navy (DON), Battelle Memorial Institute (BMI) implemented an evaluation of practices currently applied at Base Realignment and Closure (BRAC) cleanup sites involved in the characterization, classification, and disposal of soil and debris containing radium-226 (Ra-226) as low-level radiological waste (LLRW). The purpose of this evaluation was to analyze LLRW characterization and disposal practices and determine whether changes in current practices might result in a reduction in the volume of waste incorrectly classified as LLRW. The evaluation focuses on Ra-226, although other radionuclides of concern include cesium-137 (Cs-137) and strontium-90 (Sr-90).

As part of BMI activities, Argonne National Laboratory participated in the review and evaluation of LLRW generation and disposal practices at Hunters Point Shipyard (HPS), San Francisco, California. This review included meeting at HPS with Navy and contractor representatives to better understand the soil characterization and LLRW decision-making process currently in place at HPS; reviewing HPS LLRW-associated documents and datasets provided by the DON and its contractors; identifying aspects of the LLRW characterization, segregation, and disposal process most likely contributing to the volume of LLRW soil currently being generated at HPS; and suggesting modifications to current protocols to minimize the production of LLRW soil.

This report summarizes the findings of this work. The datasets analyzed in this report are from the sanitary sewer system remediation, which has been found to have, at most, minor Ra-226 contamination. However, many of the procedural modifications suggested in this report would apply equally well in areas with more significant contamination. In such areas, false detection at background levels of radionuclides of interest would be less of an issue than it is in the sanitary sewer remediation areas. Likewise, while this evaluation focuses on LLRW generation and disposal practices at HPS, the resulting recommendations would apply as appropriate to other BRAC sites.

The intent of the reviewed sewer removal and radiological survey process is to obtain radiological release of the resulting trench and to obtain unrestricted use and clearance of the excavated soil for use as clean backfill. Thus, the practices reviewed were by nature conservative with respect to the identification and segregation of impacted soil. This objective no doubt contributed to some over-identification of LLRW for soil near the cleanup level. Thus the recommendations in this report might not apply to cleanups in other areas with other objectives that use the same protocols.

Finally, many of the current characterization approaches and the criterion referenced in this report are the product of negotiations between the DON and its regulators. The recommendations contained in this report reflect a technical analysis of the issues present, and may or may not be implementable under the current negotiated regulatory environment.

2 HUNTERS POINT SHIPYARD

HPS was involved in a number of historical activities that potentially could have released radioactive material to the environment, resulting in contamination of soil above background conditions. As part of its environmental restoration program, the DON evaluates soil at impacted sites for evidence of Ra-226 and other radionuclides above negotiated standards. When such soil is identified, it is segregated for disposal at facilities outside the state of California. The following bullets summarize the current standards, their interpretation at HPS, the protocols in place to determine whether soil meets those standards, and the procedures and criterion used to segregate soil above the standards.

- Soil that exceeds background conditions by a criterion of 1 pico-Curie per gram (pCi/g) of Ra-226 is required to be segregated, removed, and disposed of outside of the state of California. Background conditions are based on a round of historical sampling that determined an average background Ra-226 activity concentration of 0.485 pCi/g, resulting in a cleanup level of 1.485 pCi/g.
- Soil that has been excavated is spread onto bermed soil “pads” to facilitate waste characterization.
- Characterization begins with a complete gross activity gamma survey of a soil pad’s exposed surface using a sodium iodide (NaI) towed-array gross activity detector which collects repeated gross activity measurements of duration on the order of seconds while being towed over the entire pad. Individual measurements are compared with a threshold of three standard deviations above the average background value calculated from on the order of 1,400 gross activity measurements collected from an uncontaminated reference area. If the threshold is exceeded at any particular location, a discrete biased sample is collected at the location.
- Characterization continues with the collection of 18 discrete systematic samples laid on a triangular grid across the pad.
- Both the biased samples and the systematic samples are analyzed on-site by a field laboratory that historically has used gamma spectroscopy focused on one of the Ra-226 characteristic energy lines for identifying and quantifying the amount of Ra-226 present.
- If any individual sample exceeds 1.485 pCi/g (average background plus 1 pCi/g), then a decision is made whether to simply excavate soil based on existing data from nearby samples, or to bound the contamination extent by collecting additional samples around the location(s) exceeding the standard.

- If a pad requires remediation, then another 18 systematic samples are collected post-remediation, and the analysis/decision-making process is repeated until a complete round of 18 samples meets the standard.

Excavated soil currently exceeds the Ra-226 criterion at a rate of about 3 to 5% by volume. While this is a relatively small fraction of total excavated soil, due to the requirement for out-of-state disposal along with transportation and management fees, it represents a significant cost.

On the basis of data provided by the DON in a spreadsheet dated November 29, 2010, 1,411 bins of contaminated soil were disposed of as LLRW totaling 22,946 cubic yards (yd³). Each of these bins was sampled, with the sample analyzed for Ra-226 by an off-site laboratory using gamma spectroscopy. The average concentration for the bins was 0.59 pCi/g, which was 0.11 pCi/g greater than the average concentration determined by previous background sampling work. Only three bin samples exceeded the 1.485-pCi/g standard. While there are various plausible explanations for why so few samples exceeded the HPS standards, these data do suggest that there is a potential for reducing the volume of soil characterized as LLRW.

3 DATA ANALYSIS

LLRW volume reduction is potentially possible via two avenues:

- By enhancing the current decision-making process to ensure that no soil is disposed of unnecessarily using the current implementation of cleanup requirements and background definitions, and
- By revisiting the way cleanup requirements are implemented and background levels are defined.

In particular, this review is concerned with the implementation of the cleanup criterion for Ra-226 in soil: 1 pCi/g above background, not to exceed 2 pCi/g. This criterion represents a difficult measurement challenge as Ra-226 is naturally present in soil at approximately 0.5 pCi/g. Adding to this challenge is the current implementation of the 1-pCi/g above background criterion as not to be exceeded in any soil volume represented by any single measurement. That is, it is being implemented on a “never-to-exceed” basis.

Incorrect identification of some small fraction of soil as exceeding the criterion is an inevitable consequence of the variability of background levels of Ra-226 activity concentrations in soil and the relatively high uncertainty of Ra-226 measurements at levels near background with available field and laboratory methods. These two factors can combine to trigger incorrect LLRW characterization when numerical measurements exceed the cleanup level as a result of variations in the background level of Ra-226 combined with measurement uncertainty, and not due to actual site contamination.

The possibility of incorrect LLRW characterization is increased by the fact that the 1-pCi/g criterion is added to the mean background concentration, rather than to a high percentile of the background distribution, such as the 95th percentile. Further, the current background level has been determined from a single set of 18 measurements in a single background reference area, which likely does not represent the full range of Ra-226 background levels in all soil types and horizons that are being remediated. Soil with average background Ra-226 levels higher than the original reference area will have higher levels of incorrect LLRW characterization.

To summarize, the challenges of applying a 1-pCi/g cleanup criterion incremental to average background conditions when background levels are around 0.5 pCi/g include the following:

- Average Ra-226 activity concentrations vary by soil type. Soil types at large facilities such as HPS can vary spatially and with depth. If the selected background reference area fails to capture all of the soil types present at the facility, it can underestimate the true average Ra-226 background level and will produce false positives for soil with an average background Ra-226 level that is higher than the reference area, or may even exceed the cleanup level as defined.

- Laboratory measurements of Ra-226 include significant measurement error at background Ra-226 activity concentrations. Potential false positives may occur if laboratory data are “noisy,” that is, if they have high relative measurement uncertainty.
- Conservative measurement assumptions may produce false positives—an example is the historical use of a single, relatively weak, gamma energy line subject to interference from naturally occurring uranium-235 (U-235) for measurement of Ra-226, combined with a point-by-point comparison of measurement results to cleanup levels.

As a result of the factors described above, background distributions will produce individual measured exceedances of the Ra-226 criterion.

The following is an analysis of several issues identified in the current LLRW program being carried out at HPS and other BRAC bases. The analysis identifies the basis of the concern, data analyses that support the concern, and possible alternative practices and approaches that might mitigate the issue. These alternatives are focused on correctly identifying soil that exceeds site radiological cleanup criteria for disposal off-site as LLRW.

3.1 EFFECTIVENESS OF THE TOWED NAI ARRAY FOR IDENTIFYING LLRW

The current soil characterization protocol requires a complete scan of the soil pad surface by NaI detectors, with the goal of identifying soil above a threshold level that potentially indicates the soil is above the Ra-226 criterion. The investigation threshold is three standard deviations above average gross activity levels collected from a reference area. Both the standard deviation and average gross activity level are based on a scan of a single background reference area located in Parcel D-2 at HPS. Biased samples are required for locations that have gross activity results exceeding the investigation threshold.

Two questions for the NaI-based scan are (1) whether it can be used to reliably infer Ra-226 activity concentrations above the Ra-226 criterion, and (2) the degree to which it produces “false positives,” that is, identifies locations as above the Ra-226 criterion when they are actually below it.

Table 1 compares Ra-226 activity concentrations in systematic and biased samples collected from remediation pads for Environmental Soil (ES) and Trench Soil Units (TU) for 2009 and 2010. The biased sampling was in response to NaI detector data. The systematic samples and the distribution of their Ra-226 activity concentrations reflect the distributions of soil in the pads. The biased sample results reflect the distributions of soil in the pads that exceeded the detector threshold. Several relevant comparisons are presented in Table 1, including the average Ra-226 activity concentration for systematic samples versus biased samples, and the percentages of each sample type that exceeded 1.485 pCi/g. Several observations are pertinent:

TABLE 1 Summary of 2009 and 2010 ES and TU Ra-226 Results for Systematic and Biased Samples

Dataset	ES 2009	ES 2010	TU 2009	TU 2010
Systematic Samples (N)	1,044	1,404	950	522
Ra-226				
Mean (pCi/g)	0.7797	0.6596	0.7241	0.5713
STDEV (pCi/g)	0.4588	0.4332	0.5364	0.5534
RSD%	58.9	65.7	74.1	96.9
Median	0.7704	0.6792	0.7043	0.5260
95th percentile	1.5359	1.2610	1.7021	1.4835
Biased Samples (N)	608	736	763	128
Ra-226				
Mean (pCi/g)	0.9001	0.7969	0.7318	0.5819
STDEV (pCi/g)	0.5339	0.4771	0.7329	0.5420
RSD%	59.3	59.9	100.2	93.1
Median	0.8755	0.7698	0.6922	0.5360
95th percentile	1.9582	1.6062	2.1464	1.4404
% Results >1.485 pCi/g				
Systematic	5.8%	2.8%	7.7%	5.0%
Biased	11.8%	7.7%	12.6%	4.7%
% Results >2 pCi/g				
Systematic	1.9%	0.3%	2.4%	1.5%
Biased	4.6%	1.8%	6.6%	2.3%

- The mean activity concentrations for biased samples are only on the order of 10% higher than systematic samples for Ra-226. This would suggest that the overall Ra-226 content of soil failing the gross gamma activity threshold was not very different from Ra-226 in the soil for the pads as a whole.
- Systematic sampling locations were not selected based on towed-array detector results. For the four groups of remediation pads represented in Table 1, the fraction of soil above 1.485 pCi/g in systematic samples ranged from approximately 3 to 8%. The percentage of biased samples exceeding 1.485 pCi/g was not much greater, suggesting that the gross gamma threshold level was not particularly successful at identifying laboratory-reported exceedances of the Ra-226 criterion.
- Only a very small percentage of biased sample results were above 1.485 pCi/g, while, for example, >50% might be expected.

This observation was confirmed in a second dataset provided by the DON. With respect to remediation decisions, the percentage of measurements exceeding cleanup levels was only slightly higher in biased samples for

144 ES units sampled from April 1, 2009, to November 11, 2010, as compared to systematic samples. (This dataset overlaps with, but is separate from, the data summarized in Table 1). For this dataset, 3.16% (82 of 2,592) of individual Ra-226 results exceeded the prevailing cleanup level of 1.485 pCi/g compared to 5.89% (101 of 1,709) for biased samples (Nov. 15, 2010, Tetra Tech e-mail). Further, the fraction of biased samples exceeding the cleanup level is quite low, given that these locations exceeded the investigation level, suggesting that the current protocol produces a large number of false positives. While these are screened out by biased sampling, sampling and analysis is itself an expensive process.

The observations above assume that the laboratory results for biased and systematic samples were themselves correct (i.e., if the detector identified an elevated area that was biased sampled and the sample result came back below 1.485 pCi/g, it was the interpretation of the towed-array detector results that was wrong, not the laboratory). Analyses in Section 3.2 of this report suggest that historically, the field laboratory results have not been a reliable measure of whether flagged biased samples exceed 1.485 pCi/g for Ra-226, given the high reported relative uncertainty in Ra-226 measurements near background. Sources of Ra-226 measurement error are discussed in Section 3.2.

NaI detectors, such as those deployed at HPS, are extremely sensitive instruments, capable of detecting small variations in background gross activity. Unfortunately, as currently configured, they are not capable of identifying the source of those variations. In addition, detector response may be affected by soil density and moisture content variations. For surface soil, background radionuclides that are present and contributing to the gross gamma activity measured by a NaI detector include Cs-137, U-238 and its decay chain (which includes naturally occurring Ra-226), thorium-232 (Th-232) and its decay chain (which includes Ra-228), and potassium (K-40). One indication of whether an increase in detector activity is associated with variations in natural background or the presence of anthropogenic contamination in the case of Ra-226 is whether slightly elevated Ra-226 levels are also associated with slightly elevated Ra-228 levels, as discussed in more detail in Section 3.4.

Figure 1 shows a NaI detector towed-array background dataset collected in May of 2010 from an approximately 500-square meter (m^2) area at HPS that includes 885 data points with a mean of 971 counts per second (cps) and a standard deviation of 45 cps. The individual measurements in Figure 1 are color-coded by their gross activity level. As is clear from this figure, the variability present has a distinct spatial pattern and normal data distribution characteristic of background activity concentrations—not the presence of contamination.

Each soil pad has approximately 1,400 measurements collected by NaI scans. Even if a pad were at background conditions and its natural variability reflected that of the reference area, there still would be a small number of measurements that would exceed the three-standard-deviation investigation level based on statistics alone. If a pad had slightly higher background conditions and/or its natural variability were greater than that observed in the reference area, the number of measurements exceeding the investigation level could increase significantly. Whether results observed to date in soil pads from the storm and sanitary sewer investigation appear to

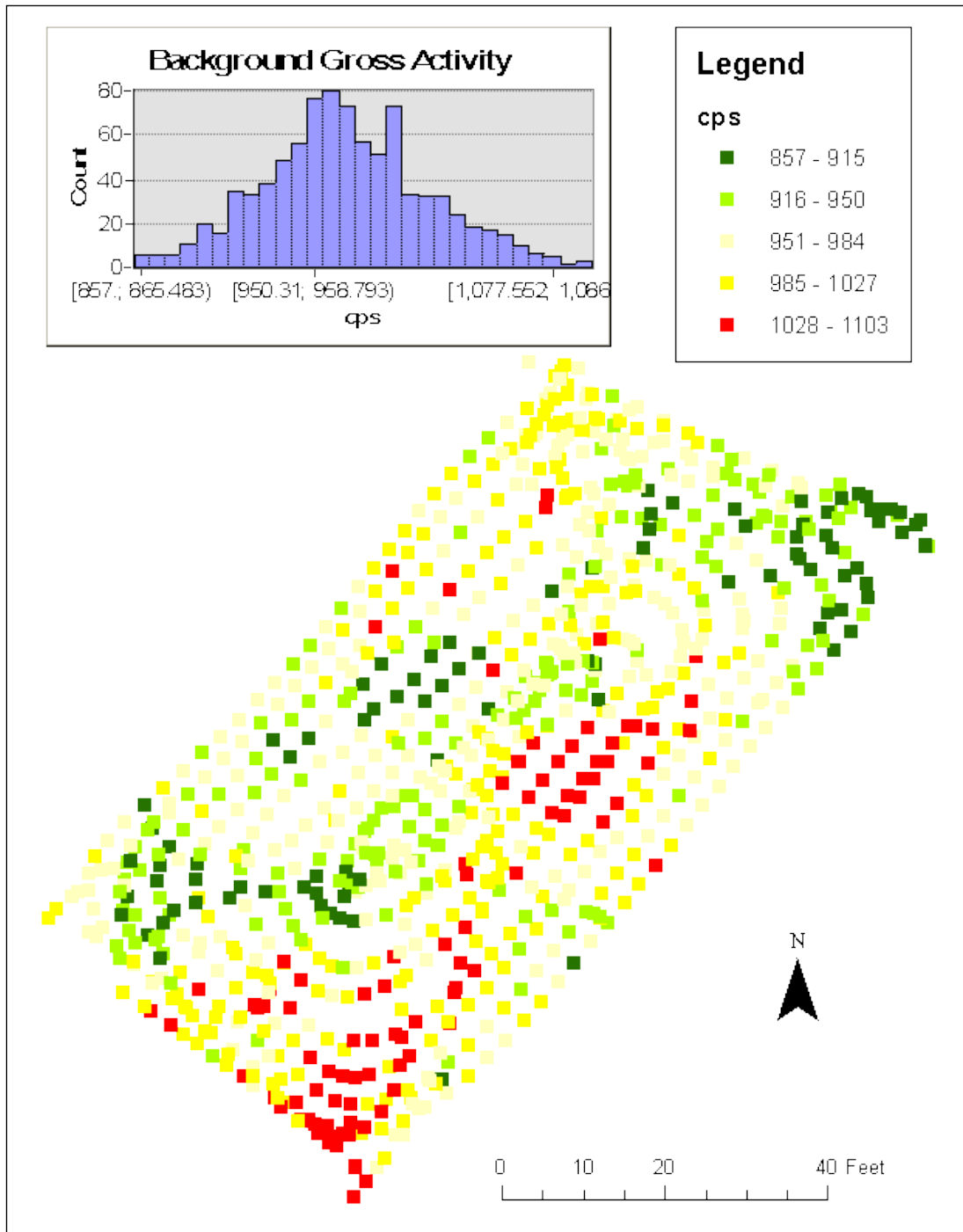


FIGURE 1 Background Area D-1 Gross Activity Data

indicate contamination or simply variations in background conditions is addressed in more detail in Section 3.4.

The conclusion is that the current NaI soil pad screening process does not appear to be a reliable method for identifying soil potentially above 1.485 pCi/g and is prone to significant false positive rates. While NaI false positives do not necessarily translate into a larger volume of LLRW, since locations are biased sampled prior to being remediated, they do increase programmatic costs because the biased sampling and analysis work can be a significant cost in itself.

The recommendations are:

- Determine the incremental detector response above average conditions that could be expected from an additional 1 pCi/g of Ra-226. While existing laboratory and scan data from pads could be used to do this, such an approach is complicated by the fact much of the current variation in average pad activity concentrations for Ra-226 appears to be natural and not anthropogenic. For example, it is likely that pads that have slightly elevated natural levels of Ra-226 would also have elevated natural levels of Ra-228. In addition, individual historical sample Ra-226 results have significant measurement error associated with them combined with a very limited range in observed activity concentrations—both of those facts would make a linear regression analysis that attempted to correlate NaI system response with variations in Ra-226 activity concentrations difficult and unlikely to be successful. An alternative approach would be to perform Monte Carlo Neutral Particle modeling for the detector system to estimate how much additional response from the system is likely to be observed from the presence of an additional 1 pCi/g of Ra-226 in soil. Once this has been determined, it could form the basis for an alternative detector threshold more likely to be associated with true Ra-226 criterion exceedances.
- Revisit the current investigation level used for selecting biased sampling locations. It likely does not reflect the true variability present in gross activity readings from background soil. One approach to developing a more inclusive investigation level would be to pool NaI results from pads where systematic sampling indicated background conditions were present to obtain a larger and likely more representative background dataset from which to derive an average background value and corresponding standard deviation. Another alternative would be to apply the three-standard-deviation rule to pads using the average gross activity for the pad and the reference area standard deviation as the basis. A third alternative would be to base both the standard deviation and average gross activity on results from the pad in question. In each alternative, the likely result would be a reduction in the number of false positives observed.

- Deploy a multichannel analyzer with the towed NaI system and perform gamma spectroscopy on the scan data. The U.S. Department of Energy Fernald site used a similar system in the late 1990s to screen soil for U-238, Th-232, and Ra-226. The system was calibrated using a specially prepared pad designed for that purpose. The advantage at HPS would be that high gross activity measurements could be screened for the presence of both Ra-228 and Ra-226 as a secondary check to minimize false positives.

3.2 EFFECTIVENESS OF ON-SITE LABORATORY FOR IDENTIFYING LLRW

HPS uses on-site gamma spectroscopy soil sample analysis for determining the activity concentration of Ra-226 in soil samples. The method relies on direct measurement of Ra-226 from its 186-kiloelectron volt (keV) emission line and produced results that had relatively high uncertainty at site soil action levels. The 186-keV line is prone to high bias from interference from U-235, which cannot be corrected because of the unavailability of other measurable gamma emissions from Ra-226. Results based on bismuth-214 (Bi-214) and lead-214 (Pb-214) progeny have lower uncertainties and less potential for high bias, but may exhibit low bias unless corrected for radon disequilibrium, either through a correction factor or through allowing sufficient time to reestablish equilibrium.

Table 2 contains Table 1 information plus Bi-214 results. The mean soil results for Bi-214 were approximately 15% lower than mean Ra-226 results for systematic and biased samples, while the Bi-214 results had standard deviations consistently about 35% lower than Ra-226. Since these nuclides are likely in secular equilibrium, this difference suggests a small bias in one or both of the measurements. The much-reduced standard deviations for Bi-214 are a consequence of lower measurement uncertainty, as indicated on gamma spectroscopy report sheets. These report sheets also indicate that Bi-214 results were typically about 10 times greater than the reported minimum detectable concentration (MDC), while Ra-226 results on the same sheets were typically near the reported MDC, with many flagged as below the MDC.

The reduced uncertainty of the Bi-214 activity concentration estimate produces a much tighter distribution of results at the background concentrations that are present in the vast majority of pad results, which produces a significantly smaller number of results above cleanup levels. As indicated at the bottom of Table 1, for a cleanup level of 1.485 pCi/g, 430 of 6,166 Ra-226 results exceeded this level, while only 75 Bi-214 results in the same set of samples exceeded this level, or 17% as many, in 2009 and 2010 ES and TU soil pads. This comparison indicates that if Bi-214 results had been used for remediation decisions, far fewer locations would have been flagged for remediation. If the average Bi-214 activity concentration in the background samples had been used instead of Ra-226 to determine a cleanup level, the number of samples exceeding the cleanup level based on Bi-214 would have been even lower.

The conclusion is that historical data indicate that the majority (as much as 80%) of historical samples with Ra-226 activity concentrations above 1.485 pCi/g likely had activity concentrations, based on Bi-214, that were in fact below that standard. The relatively high measurement uncertainty associated with measurements using the Ra-226 186-keV energy line

TABLE 2 Summary of 2009 and 2010 ES and TU Ra-226 and Bi-214 Results for Systematic and Biased Samples

Dataset	ES 2009	ES 2010	TU 2009	TU 2010
Systematic Samples (N)	1,044	1,404	950	522
Ra-226				
Mean (pCi/g)	0.7797	0.6596	0.7241	0.5713
STDEV (pCi/g)	0.4588	0.4332	0.5364	0.5534
RSD%	58.9	65.7	74.1	96.9
Median	0.7704	0.6792	0.7043	0.5260
95th percentile	1.5359	1.2610	1.7021	1.4835
Bi-214				
Mean (pCi/g)	0.6558	0.6280	0.5835	0.4826
STDEV (pCi/g)	0.2272	0.2255	0.2690	0.3096
RSD%	34.6	35.9	46.1	64.2
Median	0.6261	0.6218	0.5375	0.3906
95th percentile	1.0766	0.9580	1.0828	1.0642
Biased Samples (N)	608	736	763	128
Ra-226				
Mean (pCi/g)	0.9001	0.7969	0.7318	0.5819
STDEV (pCi/g)	0.5339	0.4771	0.7329	0.5420
RSD%	59.3	59.9	100.2	93.1
Median	0.8755	0.7698	0.6922	0.5360
95th percentile	1.9582	1.6062	2.1464	1.4404
Bi-214				
Mean (pCi/g)	0.7342	0.7177	0.6785	0.5255
STDEV (pCi/g)	0.2506	0.2701	0.3718	0.3541
RSD%	34.1	37.6	54.8	67.4
Median	0.7038	0.7047	0.6106	0.4064
95th percentile	1.1747	1.2466	1.3852	1.1498

results in a wide spread of results at background levels, including some that fell above the 1.485-pCi/g criterion due to measurement error alone.

The recommendations are:

- Change the on-site laboratory method to report Ra-226 results based on Bi-214. Structure counting times, detectors, and sample geometries such that the possibility that background soil will yield results above the criterion from measurement error alone is nil.
- Use an abbreviated in-growth period and/or a correction factor for in-growth. Site-specific correction factors can be readily determined by collecting a small set of samples, measuring them immediately on-site by using the Bi-214

energy lines, canning the samples for 30 days to allow in-growth, and then remeasuring.

- Establish a protocol that requires reanalysis of samples exceeding the 1.485-pCi/g requirement to ensure that the original observed result was not simply a product of measurement error.
- Revisit background soil sample results for Ra-226 using the improved on-site analytical methods.

3.3 HPS RA-226 BACKGROUND INTERPRETATION

The current background reference dataset likely does not fully represent the variability of background levels of Ra-226 or total soil radioactivity relevant to site soil. As a result, incorrect identification of soil for off-site disposal may be occurring. As an example, a second potential background reference area (D-1) was established and sampled systematically with 20 samples. The average Ra-226 activity concentration observed was 0.633 pCi/g with a standard deviation of 0.31 pCi/g. This is compared to an average of 0.485 and a standard deviation of 0.27 pCi/g for the currently accepted reference area (D-2). An interesting side note: because of the large degree of uncertainty associated with the individual measurements in both cases, the 95% confidence intervals around the respective means for the two background areas show a significant amount of overlap; that is, although the two average Ra-226 activity concentrations appear to be significantly different, they are not necessarily statistically significantly different.

If one had used the average background from area D-1 instead of D-2 for the data in Tables 1 and 2 (resulting in a criterion of 1.633 pCi/g rather than 1.485 pCi/g), the total number of exceedances across the ES and TU datasets would have dropped from 430 samples to 321 samples. If, in addition, the analysis had been based on the Bi-214 energy line rather than the Ra-226 energy line, there would have been only 46. This example illustrates how sensitive the application of the criterion (and consequently LLRW volumes) is to background assumptions.

For the 92 ES and TU pads characterized in 2009 and 2010, the average activity concentration for Ra-226 (as measured by the first 18 systematic samples using Bi-214) ranged between 0.15 pCi/g and 1.65 pCi/g. The 0.15-pCi/g result indicates how low average remediation pad Ra-226 activity concentrations can be—these low values would be attributable to background conditions. Likewise, one would expect to see at least as much variability in average background pad conditions above the historical background reference area average of 0.485 pCi/g, suggesting that a portion of the pads with activity concentrations above the 1.485-pCi/g cleanup level do not actually exceed the 1-pCi/g above background criterion.

One way to explore this further is to look at how average Ra-228 activity concentrations behave in these pads. There is no evidence that Ra-228 would be a contaminant of concern for this soil. In the natural environment, there typically is a fairly strong correlation between background activity concentrations for Ra-226 and Ra-228. Fortunately, the on-site gamma spectroscopy reported actinium-228 (Ac-228) activity concentrations, the daughter typically used

for estimating Ra-228 activity concentrations. The average activity concentration for Ac-228 for the 92 pads ranged from 0.16 to 1.68 pCi/g—almost exactly the same as the range for Ra-226. The conclusion is that the distribution of average Ra-226 activity concentrations in these pads mirrored that of Ra-228, which is not a contaminant of concern for the site and which exists at background levels.

Even more striking, the correlation between the average activity concentrations for these two radionuclides across the 92 pads was 0.67 ($R^2=0.45$), indicating a fairly strong positive correlation, as evidenced in the scatter plot in Figure 2. In general, relatively high Ra-226 activity concentrations were accompanied by relatively high Ra-228 activity concentrations, a relationship that would not have been the case if the higher Ra-226 activity concentrations were due to contamination.

Another piece of evidence that suggests that the preponderance of pads consisted of background soil is the NaI detector data. The NaI scanning system employed at the site has sufficient sensitivity to detect relatively subtle variations in background gross activity levels, as

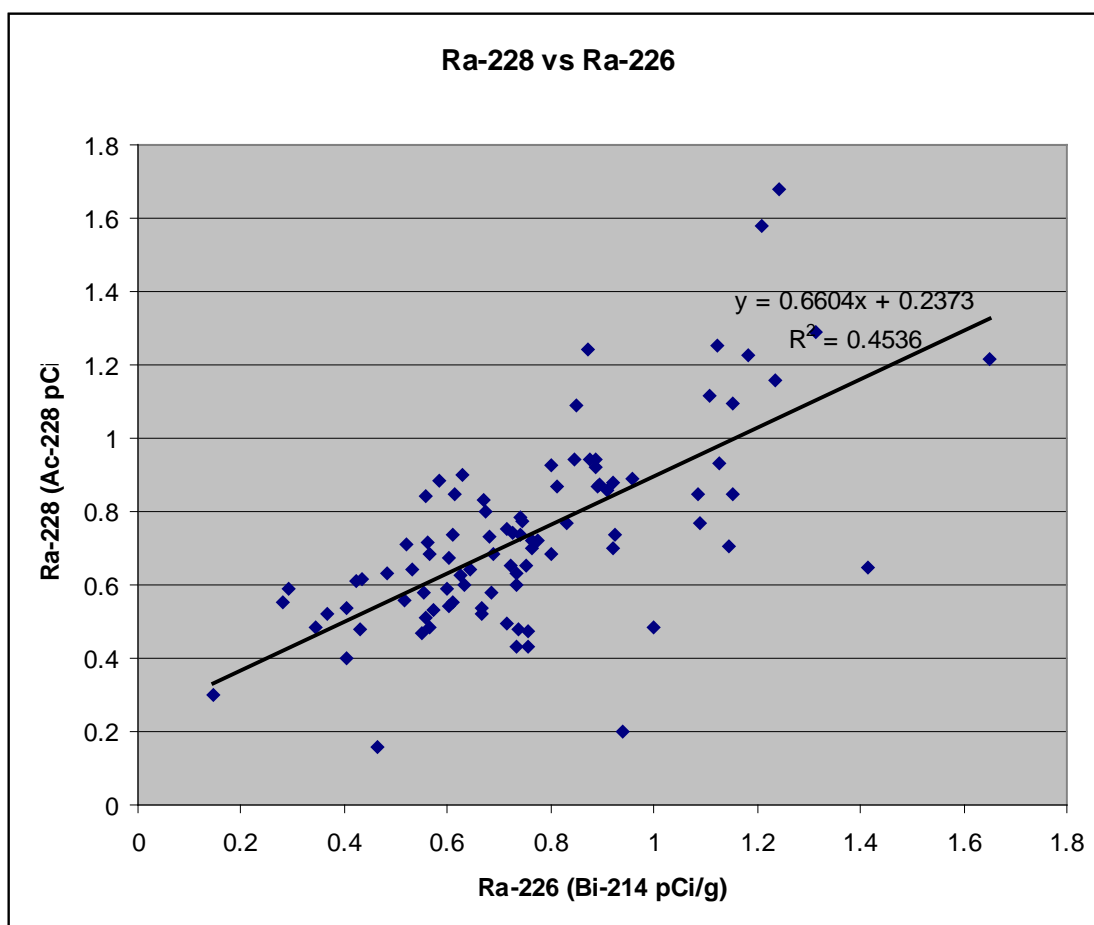


FIGURE 2 Relationship between Ra-226 and Ra-228 Average Activity Concentrations across Pads

indicated by the spatial patterns evident in Figure 1. Figure 3 maps NaI data for six soil pads that span a range of average Ra-226 activity concentrations (with Ra-226 activity concentrations estimated by using Bi-214 results). The pads are organized by increasing average Ra-226 activity concentrations. The color maps used to portray the NaI data vary from map to map as is evident from the legends. As Figure 3 demonstrates, all six soil pads showed similar types of spatial patterns in their gross activity, despite the fact that their average Ra-226 activity concentration ranged from 0.24 to 0.92 pCi/g, and despite the fact that, based on original biased and systematic sample Ra-226 results, units 199, 226, and 284 all underwent soil removal.

Two additional striking facts for these six pads are evident—the average Ra-228 activity concentration ranged over almost 1 pCi/g across the pads, indicating just how much natural average radium activity concentrations can vary. Finally, for these six pads, the connection between Ra-226 and Ra-228 was very strong, again suggesting that in reality the soil in all six pads was background soil.

On the basis of the available data, the conclusion is that the vast majority of soil screened in 2009 and 2010 that composed the reviewed dataset was likely background soil. The criterion exceedances observed were likely driven by variations in average background conditions coupled with relatively large uncertainties associated with the on-site laboratory Ra-226 measurements.

The recommendations are:

- Additional background soil sampling should be performed to more fully capture the range in natural variability present in background soil at the site. This sampling should include both surface and subsurface soil, since there often can be significant differences in background activity concentrations for naturally occurring radionuclides between surface soil and soil one or two meters deep.
- The site should consider including other indicators when determining whether a particular sample represents unacceptable contamination or simply background fluctuations. The historical data suggest that the ratio of Ra-226 to Ra-228 might be a useful indicator.
- The site should consider revising the way the 1-pCi/g above background action criterion for Ra-226 is implemented, which is complicated by the fact that the criterion is similar in magnitude to natural background variability, particularly when that variability is amplified by measurement uncertainty. It might be reasonable to add the 1-pCi/g criterion to a high percentile of the measured background distribution, such as the 95th percentile, instead of to the mean, to determine the cleanup level. Alternatively, the upper 95% confidence interval on the mean might be used as a point of departure. Either approach would reduce the fraction of false positive results due to measurement error, while still being protective. Under either approach, no soil would remain that exceeds the current 2-pCi/g upper limit.

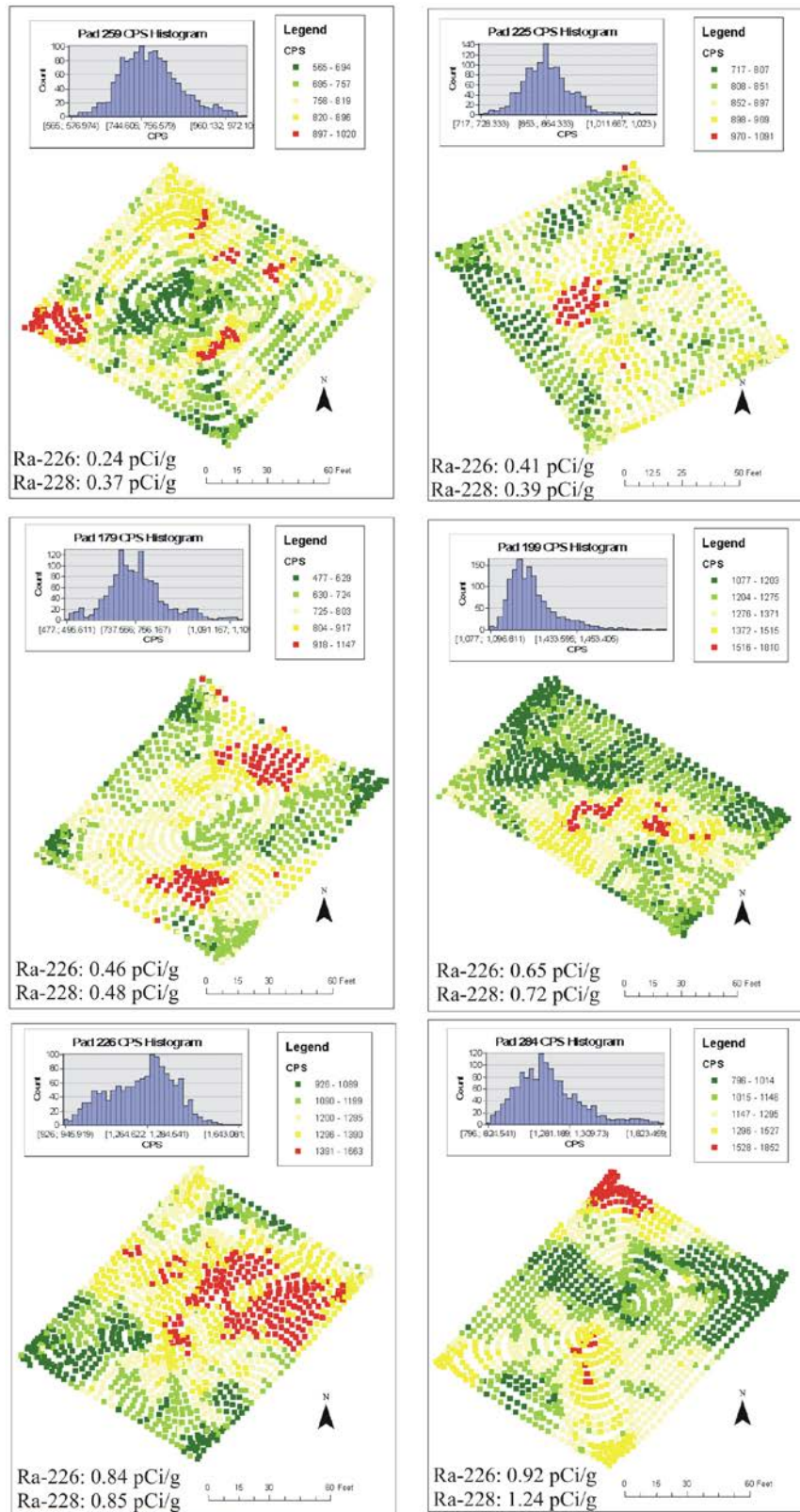


FIGURE 3 NaI Scan Results and Average Ra-226 and Ra-228 Levels for Six Soil Pads

3.4 CHARACTERIZATION PROTOCOLS

Under the current program, a second set of systematic samples is collected on pads that have been remediated (as described in Section 2) to determine that no further remediation is needed. Such an approach is not consistent with a sample-by-sample comparison approach that is currently in place for soil pads, where each sample is required to achieve the criterion. Therefore, the current practice appears to result in unnecessary sampling. While such sampling might occasionally detect locations above cleanup levels, the cost-effectiveness of such sampling—a new set of 18 systematic random samples at locations different from the initial set of systematic samples—is in question.

Given the reviewed soil pad results in 2009 and 2010 presented in Tables 1 and 2 and the discussion in Section 3.2 and 3.3, many of the results (both systematic and biased) above cleanup levels appear to be random statistical anomalies produced by large measurement error and variations in background conditions, not actual contamination that was initially inferred by data from the towed-array screening method. This conclusion includes any positive results observed in follow-up systematic sampling. Thus, the current practice of further rounds of systematic sampling is likely compounding the problem of false positive identifications and unnecessary remediation.

Ideally, under a “never-to-exceed” sampling objective, only biased sampling would be required, assuming that an effective screening method with appropriate investigation levels is in place. Biased samples would be collected at locations exceeding the scan threshold level, with the likelihood of missing contamination exceeding the criterion very low. No systematic sampling would be needed. In this context, it is unclear what role the current systematic sampling would serve. It appears to be an artifact of a sampling design intended to address a decision unit, that is, an entire pad, but one that has been superseded by a “never-to-exceed” interpretation of the cleanup criterion, which would not require systematic sampling. This overlap then results in excessive sampling overall.

Conversely, under a “never-to-exceed” sampling objective, if there is not an effective scanning methodology, then the only method for reliably identifying contamination above the criterion is systematic sampling. Currently, as discussed in Section 3.1, it is unclear based on available data whether the NaI scans are capable of reliably identifying soil that actually exceeds, but is near, the Ra-226 criterion. Historical exceedances appear to be largely driven by on-site laboratory measurement uncertainty rather than the actual presence of Ra-226 contamination above the criterion.

The conclusion is that the current requirement for repeating systematic sampling in the event that soil is remediated and removed from a soil pad serves no useful purpose.

The recommendations are:

- Eliminate repeated systematic sampling if applying remediation levels on a point-by-point, never-to-exceed, basis.

- Convert to a Multi-Agency Radiation Survey and Site Investigation Manual (MARSSIM)-based approach that addresses soil pads as decision units and employs separate but related cleanup criteria for average conditions over the soil pad as a whole and for small elevated areas. Under this scenario, the expectation would be that the elevated area criterion would be greater than the current Ra-226 criterion. For example, the prevailing 2-pCi/g Ra-226 upper limit might serve this purpose.

3.5 CONCLUSIONS AND RECOMMENDATIONS

On the basis of the data provided for HPS, the following general conclusions can be drawn:

- The bulk of the soil disposed of to date as LLRW due to Ra-226 associated with the reviewed sewer excavation datasets was likely background soil. It is noted, however, that these excavations were performed with the objective of unrestricted release of trench floors and the use of excavated soil as clean backfill, which likely contributed to this outcome.
- The technical reasons for its identification as LLRW were likely a combination of measurement error from on-site laboratory methods and a background definition that likely did not reflect the true range of background conditions on-site. Of these two contributors, on-site laboratory results measurement error was the primary driver. If these two issues are addressed, one can expect to see a significant reduction in LLRW soil produced requiring off-site disposal.
- The laboratory issue is straightforward to address and may have been improved by laboratory upgrades and protocol changes instituted by the on-site contractor in early 2011. On-site laboratory data quality should be monitored carefully going forward to ensure the original issue has been resolved. In particular, samples exceeding the criterion for Ra-226 should at minimum be reanalyzed and/or split and analyzed by an off-site laboratory to verify the initial finding. While this might not be timely enough to affect the decision-making process for the soil pad in question, it would assist in determining whether false positive problems still exist.
- At a minimum, background reference soil samples should be reanalyzed by the upgraded laboratory protocols to reestablish background values.
- The background definition issue may require renegotiation of appropriate background standards with stakeholders. One recommendation would be to observe the impact that revised laboratory procedures have on LLRW generation rates, and if there is not a significant reduction in LLRW volumes, then engage stakeholders to reevaluate background conditions. An approach to

revising the definition of background would be to expand the background reference area dataset to encompass a larger range of soil types and to look not only at the average activity concentrations observed across these soil types, but also at the degree of variability present in the dataset. To further this end, the site could reasonably propose the use of a high percentile of the background distribution, such as the 95th percentile, instead of the arithmetic mean as the point of departure for applying the 1-pCi/g Ra-226 cleanup criterion. Alternatively, the upper 95% confidence interval on the mean could be applied. Either approach would have the effect of reducing the volume of identified LLRW. Whatever alternative definition is agreed to, it should reflect the variability present in background conditions.

- The current NaI screening protocols have not been effective in identifying soil that is truly of concern, while little or no actual contamination was likely present in the reviewed data sets. The protocols currently produce high false positive rates. The source of the problem is hard to definitively identify with existing data since the historical on-site laboratory dataset has too much noise and too limited a range of activity concentrations present to effectively correlate NaI data with Ra-226 activity concentrations at the low levels encountered to date. A recommendation is that the site determine the response of the system to an additional 1 pCi/g of Ra-226 in soil. This piece of information would be invaluable in selecting a field investigation level that effectively identifies Ra-226 issues without resulting in a significant number of false positives. Alternatively, the site could consider modifying the towed-array system to measure Ra-226 and Ra-228 directly. While not necessarily contributing to LLRW volumes, the current protocols do result in increased characterization costs in the form of biased sampling requirements and the impression that contamination exists in the soil pads.
- The current sampling protocol requires redoing systematic sampling in the event that excavation is required from a pad. This technically unnecessary requirement has two undesirable results. First, it increases the overall characterization costs for the pad. Second, since historically many of the exceedances appeared to be driven by on-site laboratory data measurement error, it increased the possibility that even more locations would be identified as needing remediation, resulting in still more soil removal. The latter problem will likely have been ameliorated to some extent by on-site laboratory equipment and protocol upgrades. There is no technical reason for redoing the systematic sampling.
- Implementing MARSSIM principles in soil pad surveys and LLRW decisions should be considered. Pads would be treated as decision units. Small elevated areas would be identified via towed-array scans and confirmed with biased samples. Decision units would be evaluated against the current 1-pCi/g above background criterion implemented as a wide-area average using systematic random samples. Small, elevated areas would be evaluated against an elevated

measurement comparison criterion, which would be somewhat higher than the wide-area criterion.

- The ratio of Ra-226 to Ra-228 can be used as part of a weight-of-evidence approach in deciding as to whether soil should be classified as LLRW.
- Data from the recently installed laboratory equipment can be used to compare direct Ra-226 measurements with Bi-214 proxy results. Consider reporting Ra-226 activity based on Bi-214 results if superior performance is observed for this method. To address the radon-222 (Rn-222) ingrowth issue, consider the use of an abbreviated ingrowth period and/or the use of a conservative correction factor to reduce analysis time.